ENTROPY AND MEASURES OF MAXIMAL ENTROPY FOR AXIAL POWERS OF SUBSHIFTS

TOM MEYEROVITCH & RONNIE PAVLOV

Abstract. The notions of "limiting entropy" and "independence entropy" for one-dimensional subshifts were introduced by Louidor, Marcus, and the second author in [15]. It was also implicitly conjectured there that these two quantities are always equal. We verify this conjecture, which implies, among other things, that the limiting entropy of any one-dimensional SFT is of the form $\frac{1}{n}\log k$ for $k, n \in \mathbb{N}$. Our proof also completely characterizes the weak limits (as $d \to \infty$) of isotropic measures of maximal entropy; any such measure is a Bernoulli extension over some zero entropy factor from an explicitly defined set of measures. We also discuss connections of our results to various models and results arising in statistical mechanics.

1. Introduction

In statistical mechanics, many types of systems are known or believed to exhibit *mean field behavior* in high dimensions. The original context of this term applies to the Ising model on $\mathbb{Z}^d$, where the asymptotic behavior in high dimensions resembles the Ising model on a complete graph, also known as the "Curie-Weiss model." More loosely, the term is used to indicate that in high dimensions the distribution at different sites is approximated by *independent* interactions with a *global* "mean field." The results obtained in this paper are of a related flavor. We state and prove them in the framework of *symbolic dynamics*. Some examples illustrating the relevance to statistical mechanics are discussed in Section 5.

Symbolic dynamics is a branch of dynamics concerned primarily with the study of particular $\mathbb{Z}^d$ topological dynamical systems called subshifts. A $\mathbb{Z}^d$ subshift is defined by a finite set Σ, called an alphabet, and a (possibly infinite) set $\mathcal{F}$ of functions from Σ to finite subsets of $\mathbb{Z}^d$, which are called *configurations*. The $\mathbb{Z}^d$ subshift with alphabet Σ induced by $\mathcal{F}$, denoted by $X(\mathcal{F})$, is defined to be the set of infinite configurations in $\Sigma^{\mathbb{Z}^d}$ which do not contain any translate of any of the forbidden configurations from $\mathcal{F}$. In the special case where $\mathcal{F}$ is finite, X is called a $\mathbb{Z}^d$ *shift of finite type*, or SFT.

In this paper we obtain results describing the asymptotic behavior of high-dimensional *isotropic* $\mathbb{Z}^d$-subshifts. The adjective "isotropic" here refers to an object (e.g. subshift, measure) which exhibits the same behavior along each cardinal direction. We show that a uniformly chosen large configuration from such a subshift exhibits a nearly site-wise independent distribution when conditioned on an explicitly defined "almost-trivial" σ-algebra. The formal statement of this is deferred to Section 4, following introduction of terminology and technical definitions in Section 2.

Given a one-dimensional subshift $X = X(\mathcal{F})$, the set of possible infinite configurations in $\Sigma^{\mathbb{Z}^d}$ for which words from $\mathcal{F}$ do not occur in any row along any cardinal

direction is a d-dimensional subshift, which we call the dth *axial power* of X and denote by $X^{\otimes d}$. For instance, for $\Sigma = \{0, 1, 2\}$ and $\mathcal{F} = \{00, 11, 22\}$, $X^{\otimes d}$ contains all $\{0, 1, 2\}$-colorings of $\mathbb{Z}^d$ where adjacent sites (sites with distance 1) have distinct colors.

Our main result concerns the limiting topological entropies of $\mathbb{Z}^d$ axial powers for a fixed one-dimensional subshift, and was implicitly conjectured in [15], where key definitions and machinery were also developed, including two entropy-like quantities associated with any nonempty one-dimensional subshift X.

The first is called the *limiting entropy* of X and is denoted $h_\infty(X)$. The limiting entropy is defined as a limit of topological entropies for the sequence of d-dimensional axial powers of X. The limit here exists since the topological entropies of axial powers of X form a nonincreasing sequence; see [15] for details. The limiting entropy $h_\infty(X)$ can also be viewed as the topological entropy of a suitably defined "infinite-dimensional" axial power, which will form the foundation of our proof.

The second quantity is the *independence entropy* of X and denoted $h_{ind}(X)$. Informally, this is a measure of how much of the topological entropy of X comes from sitewise independent behavior. If a one-dimensional subshift X has alphabet Σ, then we say that a string $A_1 A_2 \ldots A_n$ of nonempty subsets of Σ is "independently legal for X" if $a_1 \ldots a_n$ appears in a point of X for every choice of $a_1 \in A_1, \ldots, a_n \in A_n$.

For example, consider the one-dimensional subshift of finite type consisting of all $x \in \{0, 1\}^{\mathbb{Z}}$ which do not contain consecutive 1s, called the $g$olden mean shift and denoted by G. Then $\{0, 1\}\{0\}\{0, 1\}$ is independently legal for G, since all four words $000, 001, 100, 101$ appear in points of G. However, $\{0, 1\}\{0, 1\}\{0\}$ is not independently legal for G, since the word 110 is illegal in G. (This is because 110 contains consecutive 1s.)

Any independently legal string of subsets for a subshift X can be thought of as a source of words appearing in points of X which are induced by sitewise independent choices. The independence entropy of X is defined as the asymptotic exponential growth rate (in n) of the maximum number of words in X so induced by a single n-letter independently X-legal string.

It was shown in [15] that $h_{ind}(X) \leq h_\infty(X)$ for all subshifts X, and implicitly conjectured that the quantities are always equal. Our main result is that this conjecture is true.

Theorem 1.1. *For any $\mathbb{Z}$ subshift X, $h_{ind}(X) = h_\infty(X)$.*

For any SFT X, it was shown in [15] that $h_{ind}(X)$ is of the form $\frac{\log k}{n}$ where $k, n \in \mathbb{N}$, and that there is an simple algorithm to compute $h_{ind}(X)$. Thus, an important consequence of our result is that the limiting entropy for subshifts of finite type can be easily computed, and is always a rational multiple of the logarithm of a natural number. This is in sharp contrast to the topological entropy of d-dimensional subshifts, for which there is no known explicit expression, even for some of the simplest nontrivial examples.

The remainder of the paper is organized as follows. Section 2 contains necessary definitions and preliminary results for our arguments and results. Section 3 contains the proof of Theorem 1.1. Section 4 contains results regarding measures of maximal entropy for our infinite-dimensional axial power X, including a uniqueness criterion. Section 5 discusses some previously existing results for specific models

and how our results fit into this framework, and Section 6 discusses some natural questions and directions for future work.

Acknowledgement: We are grateful to Brian Marcus for his encouragement and for many stimulating discussions. We would also like to thank the math department of the University of British Columbia and the Pacific institute for Mathematical Sciences for hosting a visit of the second author, when the initial part of this research took place.

2. Definitions and preliminaries

Many symbolic dynamical concepts and notation are valid for subshifts on any countable amenable group G. For our purposes, we will only consider G either equal to $\mathbb{Z}^d$ for some $d \in \mathbb{N}$, or $G = \bigoplus_{\mathbb{N}} \mathbb{Z}$, which we denote by $\mathbb{Z}^\infty$.

In other words, $\mathbb{Z}^\infty = \bigoplus_{\mathbb{N}} \mathbb{Z}$ is the countable group of infinite sequences of integers which have only finitely many non-zero terms.

Each of these groups is *amenable* because they admit a Følner sequence, which is a sequence F_n of finite subsets for which $\frac{|gF_n \triangle F_n|}{|F_n|} \underset{n\to\infty}{\longrightarrow} 0$ for all $g \in G$. For any $d \in \mathbb{N}$, $[-n,n]^d$ is a Følner sequence in $\mathbb{Z}^d$, though we will rarely need to refer to the amenability of $\mathbb{Z}^d$. For the group $\mathbb{Z}^\infty$, an example of a Følner sequence is given by $F_n = [-n,n]^n \times \{0\}^\infty$; unless explicitly stated otherwise, F_n will refer to these specific sets.

For any finite set Σ, which we call an *alphabet*, and any $d \in \mathbb{N} \cup \{\infty\}$, define the *$d$-dimensional full shift* to be the topological dynamical system given by the space $\Sigma^{\mathbb{Z}^d}$, endowed with the *shift $\mathbb{Z}^d$-action* (σ_v), defined by $(\sigma_v x)_g := x(g+v)$ for all $v, g \in \mathbb{Z}^d$ and $x \in \Sigma^{\mathbb{Z}^d}$. The full shift is always endowed with the (discrete) product topology, with respect to which each shift is a homeomorphism.

A *configuration* is defined to be any $c \in \Sigma^S$, for a finite set $S \subseteq \mathbb{Z}^d$. The set S is called the *shape* of the configuration. For any finite set of configurations $w_1, \ldots, w_n$ with disjoint shapes $S_1, \ldots, S_n$, their *concatenation* is the configuration $w_1 w_2 \cdots w_n$ with shape $\bigcup_{i=1}^n S_i$ defined by $(w_1 w_2 \cdots w_n)_{S_i} = w_i$ for $1 \le i \le n$.

A $\mathbb{Z}^d$ *subshift* is a closed subset of the $\mathbb{Z}^d$ full shift which is shift-invariant. Any $\mathbb{Z}^d$ subshift X can be defined by a (possibly infinite) set $\mathcal{F}$ of forbidden configurations in the following way:

$$X = X(\mathcal{F}) := \{x \in \Sigma^{\mathbb{Z}^d} \ : \ x_{S+n} \notin \mathcal{F} \text{ for all finite } S \subseteq \mathbb{Z}^d \text{ and } n \in \mathbb{Z}^d\}.$$

When $\mathcal{F}$ is finite, we say that $X(\mathcal{F})$ is a $\mathbb{Z}^d$ *subshift of finite type* or *SFT*. For $d = 1$, if $\mathcal{F} \subset \Sigma^{\{0,\ldots,k\}}$ we say that $X(\mathcal{F})$ is a *k-step SFT*.

For any $\mathbb{Z}^d$ subshift X, we define the *language* of X, written $\mathcal{L}(X)$, to be the set of all configurations which appear within points of X. For any finite $S \subset \mathbb{Z}^d$, denote by $\mathcal{L}(X,S) := \mathcal{L}(X) \cap \Sigma^S$ the set of configurations in $\mathcal{L}(X)$ with shape S. For any $\mathbb{Z}^d$ subshift X and $w \in \mathcal{L}(X,S)$, the set $[w] := \{x \in X \ : \ x_S = w\}$ is called the *cylinder set* of w.

Given a $\mathbb{Z}$ subshift $X \subset \Sigma^{\mathbb{Z}}$, and any $d \in \mathbb{N}$, let $X^{\otimes d} \subset \Sigma^{\mathbb{Z}^d}$ denote the $\mathbb{Z}^d$-subshift defined by

$$X^{\otimes d} = \{x \in \Sigma^{\mathbb{Z}^d} \ : \ \forall g \in \mathbb{Z}^d \ \forall i \in \{1, \ldots, d\}, \ x_{g+\mathbb{Z}e_i} \in X\},$$

where $x_{g+\mathbb{Z}e_i} \in \Sigma^{\mathbb{Z}}$ is the sequence obtained by shifting x by g and projecting it in the ith direction.

Similarly, we define $X^{\otimes\infty} \subset \Sigma^{\mathbb{Z}^\infty}$ by

$$X^{\otimes\infty} = \{x \in \Sigma^{\mathbb{Z}^\infty} \ : \ \forall g \in \mathbb{Z}^\infty \ \forall i \in \mathbb{N}, \ x_{g+\mathbb{Z}e_i} \in X\}.$$

It is clear that $X^{\otimes d}$ is always a $\mathbb{Z}^d$ subshift, since it is closed w.r.t. the product topology on $\Sigma^{\mathbb{Z}^d}$, and is invariant with respect to the shift $\mathbb{Z}^d$-action on $\Sigma^{\mathbb{Z}^d}$.

We start with a simple lemma showing that for a fixed d-dimensional shape S, the set of legal configurations with shape S in $X^{\otimes d'}$ is the same for any $d' > d$.

Lemma 2.1. *For any $\mathbb{Z}$ subshift X, any $d \in \mathbb{N}$, and any finite $S \subseteq \mathbb{Z}^d$, $\mathcal{L}(X^{\otimes d}, S) = \mathcal{L}(X^{\otimes\infty}, S)$.*

Proof. The inclusion $\mathcal{L}(X^{\otimes\infty}, S) \subseteq \mathcal{L}(X^{\otimes d}, S)$ is trivial since the projection $x_{\mathbb{Z}^d}$ is clearly in $X^{\otimes d}$ for any $x \in X^{\otimes\infty}$. It therefore suffices to show the reverse inclusion.

Consider any $w \in \mathcal{L}(X^{\otimes d}, S)$. By definition, there exists $x \in X^{\otimes d}$ so that $x_S = w$. We will use x to construct a point of $X^{\otimes\infty}$. Define $x' \in \Sigma^{\mathbb{Z}^\infty}$ by $x'(g) := x(g_1, g_2, \ldots, g_{d-1}, \sum_{i=d}^{\infty} g_i)$. Clearly $x'_{\mathbb{Z}^d} = x$, so $x'_S = w$. Also, for any $g \in \mathbb{Z}^\infty$ and $i \in \mathbb{N}$, the row $x'_{g+\mathbb{Z}e_i}$ is a row of x; either in the x_i-direction if $i < d$ or in the x_d-direction if $i \geq d$. Since $x \in X^{\otimes d}$, all such rows are in X, and so $x' \in X^{\otimes\infty}$. We have then shown that $w \in \mathcal{L}(X^{\otimes\infty}, S)$, completing the proof.

□

The *topological entropy* of a $\mathbb{Z}^d$ subshift X $(d \in \mathbb{N} \cup \{\infty\})$ is defined as

$$h(X) = \lim_{k\to\infty} \frac{1}{|F_k|} \log |\mathcal{L}(X, F_k)|,$$

where F_k is a Følner sequence for $\mathbb{Z}^d$.

We define the *limiting entropy* of a $\mathbb{Z}$ subshift X as

$$h_\infty(X) := \lim_{d\to\infty} h(X^{\otimes d}).$$

The limit exists because $h(X^{\otimes d})$ is nonincreasing in d; see [15] for a proof.

The next few definitions involve measures on $\mathbb{Z}^d$ subshifts ($d \in \mathbb{N} \cup \{\infty\}$. All such measures will be taken to be Borel probability measures with respect to the product topology. We say that μ is *shift-invariant* if $\sigma_v \mu = \mu$ for all $v \in \mathbb{Z}^d$.

For a shift-invariant measure μ on X, the *measure-theoretic entropy* of μ is

$$h(\mu) := \lim_{k\to\infty} \frac{1}{|F_k|} H_\mu(F_k),$$

where

$$H_\mu(F_k) := -\int \log \mu([x_{F_k}]) d\mu(x)$$

is the Shannon entropy of the discrete random variable x_{F_k}.

The following fundamental theorem of entropy theory generalizes to any countable amenable group, and beyond. However, our statement deals only with the amenable groups $X^{\otimes d}$.

Theorem (Variational principle). *For any $d \in \mathbb{N} \cup \{\infty\}$ and $\mathbb{Z}^d$ subshift X,*

$$h(X) = \sup_\mu h(\mu),$$

where the supremum on the right-hand side is over all shift-invariant measures on X.

For a proof of the variational principle in the general amenable group setting, see [17]. For an elegant $\mathbb{Z}^d$-proof ($d < \infty$), see [16].

Since the shift $\mathbb{Z}^d$-action on the d-dimensional full shift is always *expansive*, the function $\mu \mapsto h(\mu)$ is *upper semi-continuous* for measures on any $\mathbb{Z}^d$ subshift. It follows that the supremum on the right hand side is attained by at least one measure μ. For proofs and details see [1], for instance.

A measure μ for which $h(X) = h(\mu)$ is called a *measure of maximal entropy* on X.

We also need the notion of conditional entropy. For any finite $F \subset \mathbb{Z}^d$, $d \in \mathbb{N} \cup \{\infty\}$, any measure μ on the d-dimensional full shift, and any σ-algebra $\mathcal{A}$ which is measurable with respect to μ, we define

$$H_\mu(F \mid \mathcal{A}) := -\int \log \mu([x_F] \mid \mathcal{A})\ d\mu(x),$$

and correspondingly define the *conditional entropy of* μ *with respect to* $\mathcal{A}$ as

$$h(\mu \mid \mathcal{A}) := \lim_{k\to\infty} \frac{1}{|F_k|} H_\mu(F_k \mid \mathcal{A}).$$

A standard application of the definition of conditional entropy shows that we can also write

$$h(\mu \mid \mathcal{A}) = \lim_{k\to\infty} \int \frac{1}{|F_k|} H_{\mu|\mathcal{A}}(F_k)\ d\mu(x),$$

where $\mu \mid \mathcal{A}$ denotes the conditional measure of μ given $\mathcal{A}$.

We will typically consider the case where $\mathcal{A}$ is the image of the underlying σ-algebra for the measure space under a μ-measurable factor map π; in this case, we will write the factor π in lieu of $\mathcal{A}$.

The following two well-known results relate non-conditional and conditional entropies.

Theorem. (Rokhlin-Abramov formula [23]**)** *For any shift-invariant measure μ on a $\mathbb{Z}^d$ subshift X ($d \in \mathbb{N} \cup \{\infty\}$), and any μ-measurable factor map π on X,*

$$h(\mu \mid \pi) = h(\mu) - h(\pi(\mu)).$$

Theorem 2.2. (See, for instance, p. 318 of [8]) *For any $d \in \mathbb{N} \cup \{\infty\}$, denote by P_d the lexicographic past of 0 in $\mathbb{Z}^d$, i.e. the set of all $g \in \mathbb{Z}^d$ which have at least one nonzero coordinate, the first of which is negative. Then for any shift-invariant measure μ on $\Sigma^{\mathbb{Z}^d}$, $h(\mu) = h_\mu(\{0\} \mid \pi_{P_d})$.*

(Here and elsewhere, for a set $S \subseteq \mathbb{Z}^d$, π_S represents the projection map $x \mapsto x_S$.)

The proof of Lemma 2.2 in [8] formally applies only to the case $d < \infty$, but a simple limiting argument yields the $d = \infty$ case.

The above definitions and statements about entropy are all independent of the particular choice of Følner sequence F_n. For general references on the entropy theory for amenable groups, see for example [17, 19].

We will naturally identify $\mathbb{Z}^{d_1}$ as a subgroup of $\mathbb{Z}^{d_2}$ whenever $d_1 < d_2 \leq \infty$ via the embedding $(v_1, \ldots, v_{d_1}) \mapsto (v_1, \ldots, v_{d_1}, 0, \ldots)$, and say that a sequence of measures $\mu_n \in \mathbb{P}(\Sigma^{\mathbb{Z}^n})$ converges in the weak-$*$ topology to a measure μ in $\mathbb{P}(\Sigma^{\mathbb{Z}^\infty})$ if $\mu_n([w_F]) \to \mu([w_F])$ for all finite $F \subset \mathbb{Z}^\infty$ and all $w_F \in \Sigma^F$.

The relevance of $X^{\otimes\infty}$ to the problem of limiting entropy is manifested by the following lemmas:

Lemma 2.3.
$$h(X^{\otimes\infty}) = h_\infty(X).$$

Proof. By definition, we have
$$h_\infty(X) = \lim_{d\to\infty} h(X^{\otimes d}).$$
We have
$$h(X^{\otimes d}) = \lim_{N\to\infty} \frac{1}{|[-N,N]^d|} |\mathcal{L}(X^{\otimes d}, [-N,N]^d)|.$$
Thus, we can find an increasing sequence of integers $(N_1, N_2, \ldots)$, so that
$$h_\infty(X) = \lim_{d\to\infty} \frac{1}{|[-N_d,N_d]^d|} |\mathcal{L}(X^{\otimes d}, [-N_d,N_d]^d)|.$$
As $F^{(d)}_{N_d}$ is a Følner sequence in $\mathbb{Z}^\infty$, we also have
$$h(X^{\otimes\infty}) = \lim_{d\to\infty} \frac{1}{|[-N_d,N_d]^d|} |\mathcal{L}(X^{\otimes\infty}, [-N_d,N_d]^d)|.$$
Since $\mathcal{L}(X^{\otimes\infty}, [-N_d,N_d]^d) = \mathcal{L}(X^{\otimes d}, [-N_d,N_d]^d)$ by Lemma 2.1, it follows that $h_\infty(X) = h(X^{\otimes\infty})$. □

Lemma 2.4. *Let $\mu_1, \mu_2, \ldots, \mu_n, \ldots$ be a sequence of measures where μ_n is a measure on $X^{\otimes n}$.*

(1) *If $\mu_n \to \mu$ in the weak-$*$ topology, where μ is a measure on $\Sigma^{\mathbb{Z}^\infty}$, then $\mu(X^{\otimes\infty}) = 1$.*
(2) *If, in addition, every μ_n is a measure of maximal entropy of $X^{\otimes n}$, then μ is a measure of maximal entropy of $X^{\otimes\infty}$.*

Proof. To prove (1), choose any sequence μ_n which approaches a weak-$*$ limit μ. Choose any finite $F \subset \mathbb{Z}^\infty$. For some sufficiently large d_0, $F \subset \mathbb{Z}^{d_0}$, by definition of $\mathbb{Z}^\infty$. By Lemma 2.1, $\mathcal{L}(X^{\otimes d_0}, F) = \mathcal{L}(X^{\otimes d}, F)$ for all $d_0 \le d \le \infty$. Thus, for all $d \ge d_0$, $\mu_d(x_F \in \mathcal{L}(X^{\otimes d}, F)) = 1$. It follows that $\mu(\{x_F \in \mathcal{L}(X^{\otimes\infty}, F)\}) = 1$. Thus,
$$\mu(X^{\otimes\infty}) = \mu\left(\bigcap_{F\subset\mathbb{Z}^\infty} \{x_F \in \mathcal{L}(X^{\otimes\infty}, F)\}\right) = 1,$$
where the intersection is over all finite $F \subset \mathbb{Z}^\infty$. This completes the proof of (1).

Now, (2) follows by combining (1), the relation $h(X^{\otimes\infty}) = h_\infty(X)$ from Lemma 2.3, and upper semi-continuity of measure-theoretic entropy. □

2.1. **Multi-choice subshifts and independence entropy.** We now recall some definitions and results from [15]. Let $\hat{\Sigma}$ denote the set of non-empty subsets of Σ. Let X be a $\mathbb{Z}^d$ subshift over Σ, where $d \in \mathbb{N}\cup\{\infty\}$. The *multi-choice shift* $\widehat{X} \subset \hat{\Sigma}^{\mathbb{Z}^d}$ is the set
$$\widehat{X} = \{\hat{x} \in \hat{\Sigma}^{\mathbb{Z}^d} \ : \ \hat{x} \subset X\},$$
where $\hat{x} \in \hat{\Sigma}^{\mathbb{Z}^d}$ is naturally interpreted as a subset of X, obtained by Cartesian products.

The *independence score* of a configuration $\hat{w} \in \hat{\Sigma}^F$ with shape F is defined by
$$S(\hat{w}) = \frac{1}{|F|} \sum_{n\in F} \log |\hat{x}_n|.$$

We analogously define the independence score for $\hat{x} \in \hat{\Sigma}^{\mathbb{Z}^d}$ by

$$S(\hat{x}) = \limsup_{n\to\infty} S(\hat{x}|_{F_n}).$$

Finally, we define the independence score of any shift-invariant measure $\hat{\mu}$ on $\widehat{X}$ as

$$S(\hat{\mu}) = \int S(\hat{x}) d\hat{\mu}(\hat{x}).$$

Observe that for any shift-invariant measure $\hat{\mu}$ and finite $F \subset \mathbb{Z}^d$, $S(\mu) = \int S(\hat{x}_F) d\hat{\mu}(\hat{x})$. In particular, since $\hat{x} \to S(\hat{x})$ is a function which is invariant under shifts, it follows that for an ergodic measure $\hat{\mu}$, $S(\hat{x}) = S(\hat{\mu})$ $\hat{\mu}$-almost everywhere. Also, by the pointwise ergodic theorem, the lim sup in the definition of $S(\hat{x})$ is actually a limit μ-almost surely.

Following [15], we define the *independence entropy* $h_{ind}(X)$ of a $\mathbb{Z}$ subshift X as

$$h_{ind}(X) = \lim_{n\to\infty} \left(\sup\{S(\hat{w}) \ : \ \hat{w} \in \mathcal{L}(\hat{X}, F_n)\} \right).$$

See Section 4 of [15] for details on how existence of the limit follows from sub-additivity.

The following lemma states convenient equivalent definitions of the independence entropy:

Lemma 2.5. *Let X be a $\mathbb{Z}^d$ subshift. The following are equivalent definitions of the independence entropy h_{ind}:*

(1) $h_{ind}(X) = \sup\{S(\hat{x}) \ : \ \hat{x} \in \hat{X}\}$.
(2) $h_{ind}(X)$ *is equal to* $\sup_{\hat{\mu}} S(\hat{\mu})$ *where the supremum is over shift-invariant measures on* $\hat{X}$.

Proof. Clearly, for any μ, $S(\mu) \leq \sup\{S(\hat{x}) \ : \ \hat{x} \in \hat{X}\}$. Since $\hat{x} \to S(\hat{x})$ is a shift-invariant function, it is constant for any ergodic μ. In fact, for ergodic μ, $S(\hat{x}) = S(\mu)$ μ-almost surely.

For any given $\epsilon > 0$, there are infinitely many n's and $\hat{w}_n \in \mathcal{L}(\hat{X}, F_n)$ with $S(\hat{w}) \geq h_{ind}(X) - \epsilon$. Since $S(\hat{w})$ is an average of $S(w')$ for subwords w' of smaller length, we can assume $\hat{w}_n$ is a prefix of $\hat{w}_{n+1}$. By compactness of $\hat{X}$, there is $\hat{x} \in \hat{X}$ with $S(\hat{x}) \geq h_{ind}(X) - \epsilon$. It follows that the orbit closure of $\hat{x}$ supports an invariant measure μ with $S(\mu) \geq h_{ind}(X) - \epsilon$ (see [6] for a classical presentation of this correspondence principle).

□

2.2. **Exchangeability and de Finetti's theorem.** One of the main tools in the proof of Theorem 1.1 is de Finetti's Theorem, which we review here for completeness.

Let $\mathbb{P} := Perm(\mathbb{N})$ denote the group of finite permutations of the positive integers, i.e. permutations of $\mathbb{N}$ which fix all but finitely many integers. $\mathbb{P}$ is a countable, amenable group.

Definition 1. *A sequence (X_n) of random variables is called **exchangeable** if for any n and n-tuples of distinct integers $i_1, i_2, \ldots, i_n$ and $j_1, j_2, \ldots, j_n$, the joint distributions of $X_{i_1}, \ldots, X_{i_n}$ and $X_{j_1}, \ldots, X_{j_n}$ are the same. Equivalently, (X_n) is exchangeable if all joint distributions are invariant under the action of $\mathbb{P}$ on (X_n) by permutation of indices.*

The simplest examples of exchangeable sequences of random variables are i.i.d. or Bernoulli sequences, which are clearly exchangeable. However, not every exchangeable sequence is Bernoulli. For instance, one can define the X_i so that either they must all be equal to 0 (say with probability 0.3) or all must be equal to 1 (say with probability 0.7). The reader may check that this sequence is exchangeable, but clearly it is not Bernoulli. However, it is a mixture of the two (trivial) Bernoulli sequences which are a.s. constant with value 0 and a.s. constant with value 1 respectively. In fact, this is not an anomaly: de Finetti's theorem states that all exchangeable sequences are mixtures of i.i.d. distributions.

Theorem (de Finetti's theorem). *Any exchangeable sequence (Y_n) of random variables each taking values in a finite set Ω is a mixture of identically distributed random variables. In other words, there is a measure θ on the simplex of probability measures on Ω such that*

$$P\left(\bigcap_{k=1}^{N}\{Y_k = a_k\}\right) = \int \prod_{k=1}^{N} p(a_k)d\theta(p),$$

for any $a_1, \dots a_N \in \Omega$.

For our purposes, de Finetti's theorem for finite-valued exchangeable sequence is sufficient. We note that there is a stronger version, due to Hewitt and Savage ([10]), which applies to random variables taking values in a Borel measurable space.

In particular, for exchangeable random variables, the *exchangeable* σ-algebra coincides with the *tail* σ-algebra: Any measurable function of an exchangeable sequence which is invariant under finite permutations of the variables is measurable with respect to the tail. In the case of exchangeable variables taking values in a finite set Ω, the particular i.i.d. distribution in the mixture can be recovered by observing the *empirical distributions* $\overline{\mu}_a(X_1, \dots, X_n, \dots) := \lim_{n\to\infty} \frac{1}{n}|\{i \leq n \ : \ X_i = a\}|$, $a \in \Omega$.

2.3. **Allowable local perturbations.** A useful notion in the proof and statement of our results is the notion of "allowable local perturbations" for a point in a $\mathbb{Z}^d$ subshift.

For any subshift $X \subset \Sigma^{\mathbb{Z}^d}$, $d \in \mathbb{N} \cup \{\infty\}$, shift-invariant measure μ on X, $x \in X$, and $g \in \mathbb{Z}^d$, we say that a letter a is a *μ-allowable local perturbation of x at g* if the conditional probability

$$p_{x,g}(a) := \mu(\{z \in X : z_g = a\} \mid \{z \ : \ z_{\{g\}^c} = x_{\{g\}^c}\})$$

is greater than 0.

For any shift-invariant measure μ on X, define a map $\pi_\mu : X \to \hat{\Sigma}^{\mathbb{Z}^d}$ by

$$\pi_\mu(x)_g = \{a \in \Sigma \ : \ p_{x,g}(a) > 0\}.$$

This is a measurable factor map to $\hat{\Sigma}^{\mathbb{Z}^d}$; μ-almost surely, $(\pi_\mu(x))_g \neq \varnothing$ for all $g \in \mathbb{Z}^d$, since by definition $\mu(\{x_g \in \pi_\mu(x)_g\}) = 1$.

3. Proof of theorem 1.1

The group $\mathbb{P}$ acts on $\mathbb{Z}^\infty$ in a natural way by permuting coordinates: for $\rho \in \mathbb{P}$ and $g \in \mathbb{Z}^\infty$, $(\rho(g))_i := g_{\rho(i)}$. Through the action of $\mathbb{P}$ on $\mathbb{Z}^\infty$, $\mathbb{P}$ also acts on $X^{\otimes\infty}$.

We will consider measures of maximal entropy on $X^{\otimes\infty}$ which are in addition invariant with respect to the action of $\mathbb{P}$ on $X^{\otimes\infty}$. The existence of such measures

follows from amenability of $\mathbb{P}$, upper semi-continuity of measure-theoretic entropy, and the fact that the action of $\mathbb{P}$ on $X^{\otimes\infty}$ preserves measure-theoretic entropy.

To be specific, choose any shift-invariant measure ν on $X^{\otimes\infty}$, and take any weak-$*$ limit point μ of the sequence $\frac{1}{|P_n|}\sum_{\rho\in P_n}\rho\circ g(\nu)$, where the sum is over $g\in F_n$ and $\rho\in P_n$, where $P_n\subset\mathbb{P}$ is the set of permutations which fix all integers greater than n.

Clearly, $h(\nu)=h(\rho\nu)$ for any permutation $\rho\in\mathbb{P}$. Since $\nu\mapsto h(\nu)$ is an affine and upper semi-continuous function, it follows that μ is a measure of maximal entropy whenever ν is a measure of maximal entropy.

Lemma 3.1. *For any $\mathbb{P}$-invariant measure ν on $X^{\otimes\infty}$ and any finite $F\subset\mathbb{Z}^\infty\setminus\{0\}$, x_0 and x_F are conditionally ν-independent given $x_{F^c\setminus\{0\}}$.*

Proof. Define $m=\max_{g\in F}\max\{k\in\mathbb{N}\ :\ g_k\neq 0\}$, and define $\tau:\mathbb{Z}^\infty\to\mathbb{Z}^\infty$ by $(g_1,g_2,\dots)\mapsto(0,g_1,g_2,\dots)$. It follows from $\mathbb{P}$-invariance of ν that

$$x_F,x_{\tau^m F},x_{\tau^{2m}F},\dots$$

is an exchangeable sequence of random variables. For any $a\in\Sigma$ and sequence (C_n) of elements of Σ^F, define

$$f_a(C_0,C_1,C_2,\dots):=\nu(x_0=a\mid x_F=C_0,x_{\tau^m F}=C_1,x_{\tau^{2m}F}=C_2,\dots).$$

f_a is a function of $C_0,C_1,\dots$ which is invariant under all finite permutations. It follows from de Finetti's theorem that it is measurable with respect to the tail σ-algebra. This proves that x_0 is conditionally independent from x_F given $\{x_{\tau^{km}F}\ :\ k\geq 1\}$, and by our choice of m, $\tau^{km}F\subset F^c$ for all k.

□

Lemma 3.2. *Let μ be a measure on $X^{\otimes\infty}$ which is both shift-invariant and $\mathbb{P}$-invariant. Then for a set of $x\in X^{\otimes\infty}$ of full μ-measure, any $y\in\Sigma^{\mathbb{Z}^\infty}$ obtained by a finite number of μ-allowable local-perturbations is in $X^{\otimes\infty}$. Also, for any such x, $\pi_\mu(x)\in\widehat{X^{\otimes\infty}}$.*

Proof. By compactness of $X^{\otimes\infty}$, it is sufficient to show that for μ-a.e. $x\in X^{\otimes\infty}$, any finite $F\subset\mathbb{Z}^\infty$, and any configuration $y_F\in\Sigma^F$ with $y_g\in\pi_\mu(x)_g$ for all $g\in F$, there exists $z\in X^{\otimes\infty}$ with $z_F=y_F$.

It will suffice to show that for μ-a.e. $x\in X^{\otimes\infty}$ and for any choices $y_g\in\pi_\mu(x)_g$ for all $g\in F$,

$$\mu(\{z\in X^{\otimes\infty}:z_F=y_F\mid z_{F^c}=x_{F^c}\})=\prod_{g\in F}p_{x,g}(y_g).\tag{1}$$

This is because integrating (1) with respect to μ shows that $\mu([y_F])>0$, and since μ is a measure on $X^{\otimes\infty}$, this would show that $y_F\in\mathcal{L}(X^{\otimes\infty})$, completing the proof as explained above.

We prove (1) by induction on $n=|F|$. For $n=1$, (1) is just the definition of $p_{x,g}$. For the inductive step, assume that $|F|=n$ and that (1) is true for all sets of cardinality less than n.

Choose some $g\in F$, and let $F_1:=F\setminus\{g\}$. For μ-a.e. $x\in X^{\otimes\infty}$, the definition of conditional probability gives

$$\mu(\{z\in X^{\otimes\infty}:z_F=y_F\mid z_{F^c}=x_{F^c}\})=\mu(\{z\in X^{\otimes\infty}:z_{F_1}=y_{F_1}\mid z_{F_1^c}=y_{F_1^c}\})p_{x,g}(y_g).$$

The inductive hypothesis on F_1 implies that

$$\mu(\{z \in X^{\otimes\infty} : z_{F_1} = y_{F_1} \mid z_{F_1^c} = y_{F_1^c}\}) = \prod_{h \in F_1} p_{y,h}(y_h),$$

so showing (1) for F has been reduced to showing that $p_{x,g}(y_g) = p_{y,g}(y_g)$.

To see this, note that Lemma 3.1 implies that μ-almost surely with respect to x, for any finite $F \subset \mathbb{Z}^\infty \setminus \{g\}$,

$$p_{x,g}(y_g) = \mu(\{z \in X^{\otimes\infty} : z_g = y_g \mid z_{\{g\}^c} = x_{\{g\}^c}\}) = \mu(\{z \in X^{\otimes\infty} : z_g = y_g \mid z_{F^c} = x_{F^c}\}).$$

In particular,

$$p_{x,g}(y_g) = p_{y,g}(y_g)$$

since x and y are identical off the finite set F, and we are done. $\square$

In other words, for μ as above, the measure-theoretic factor map π_μ maps to $\widehat{X^{\otimes\infty}}$, μ-almost surely. We denote by $\hat{\mu}$ the measure $\pi_\mu(\mu)$ on $\widehat{X^{\otimes\infty}}$, which is just the pushforward of μ under π_μ.

Our next step is to characterize the structure of $\mathbb{P}$-invariant measures of maximal entropy on $X^{\otimes\infty}$.

Lemma 3.3. *For any $\mathbb{P}$-invariant measure of maximal entropy μ on $X^{\otimes\infty}$,*

$$\mu = \int_{\widehat{X^{\otimes\infty}}} \mu_{\hat{x}} \, d\hat{\mu}(\hat{x}), \tag{2}$$

where for any $\hat{x} \in \widehat{X^{\otimes\infty}}$, $\mu_{\hat{x}}$ denotes the independent product of the uniform measures on $\hat{x}_g$ over all $g \in \mathbb{Z}^\infty$.

Proof. The conditional entropy of μ given π_μ satisfies

$$h(\mu \mid \pi_\mu) \ = \int_{\widehat{X^{\otimes\infty}}} H_\mu(\{0\} \mid \pi_\mu(x)) \, d\hat{\mu} \le \int_{\widehat{X^{\otimes\infty}}} \log |\hat{x}_0| \, d\hat{\mu}(\hat{x}),$$

with equality holding in the second inequality iff the conditional distribution of x_0 under μ given $\pi_\mu(x)$ is μ-a.s. uniform over the finite set $\pi_\mu(x)_0$.

Thus, if μ does not satisfy the formula (2), we could define a measure $\nu = \int_{\widehat{X^{\otimes\infty}}} \mu_{\hat{x}} \, d\hat{\mu}(\hat{x})$ for which $h(\nu) > h(\mu)$, which would contradict the fact that μ is a measure of maximal entropy. $\square$

Lemma 3.4. *Let μ be a $\mathbb{P}$-invariant measure of maximal entropy on $X^{\otimes\infty}$. The measure-theoretic entropy of the measure $\hat{\mu}$ with respect to the $\mathbb{Z}^\infty$-action by shifts is zero:*

$$h(\hat{\mu}) = 0.$$

Proof. Suppose for a contradiction that $h := h(\hat{\mu}) > 0$. First, we note that by definition, for any $g \in \mathbb{Z}^\infty$, $\pi_\mu(x)_g$ is μ-a.s. uniquely determined by $x_{\{g\}^c}$ as the set $\{a \in \Sigma \ : \ p_{x,g}(a) > 0\}$. But by Lemma 3.3, μ-a.s., $\pi_\mu(x)_g$ is independent of $x_{\{g\}^c}$ given $\pi_\mu(x)_{\{g\}^c}$. It is clear that if two random variables are independent and the first is a function of the second, then the first must be constant. Therefore, $\pi_\mu(x)_g$ must be conditionally constant given $\pi_\mu(x)_{\{g\}^c}$. In other words, μ-a.s., $\pi_\mu(x)_{\{g\}^c}$ uniquely determines $\pi_\mu(x)_g$.

This means that there exists N so that $\pi_\mu(x)_0$ is determined by $\pi_\mu(x)_{F_N\setminus\{0\}}$ with μ-probability $1-\delta$, for $\delta < \frac{h}{3\log|\Sigma|}$. Choose d large enough that $\frac{1}{|F_d|}H_{\hat\mu}(F_d) \le h+\epsilon$, where $\epsilon < \frac{h}{3|F_N|}$. We now decompose $H_{\hat\mu}(F_N \times F_d)$ as a sum of conditional entropies:

$$H_{\hat\mu}(F_N \times F_d) = H_{\hat\mu}((F_N \setminus \{0\}) \times F_d) + H_{\hat\mu}(\{0\} \times F_d \mid \pi_{(F_N\setminus\{0\})\times F_d})$$

$$\le (|F_N|-1)|F_d|(h+\epsilon) + |F_d|\delta\log|\Sigma|$$

$$= |F_N||F_d|\left[h - \frac{h}{|F_N|} + \epsilon - \frac{\epsilon}{|F_N|} + \frac{\delta\log|\Sigma|}{|F_N|}\right] \le |F_N||F_d|\left[h - \frac{h}{|F_N|} + \epsilon + \frac{\delta\log|\Sigma|}{|F_N|}\right]$$

$$< |F_N||F_d|\left[h - \frac{h}{|F_N|} + \frac{h}{3|F_N|} + \frac{h}{3|F_N|}\right] = |F_N||F_d|h\left(1 - \frac{1}{3|F_N|}\right).$$

From the general theory of entropy for amenable groups (as in [17], p. 59), for any finite set $P \subset \mathbb{Z}^\infty$, $\frac{1}{|P|}H_{\hat\mu}(P) \ge h$, and so we have a contradiction.

□

Proof of Theorem 1.1. Let μ be a $\mathbb{P}$-invariant measure of maximal entropy on $X^{\otimes\infty}$. By Lemma 3.2, π_μ is a measure-theoretic factor from the $\mathbb{Z}^\infty$-system $(X^{\otimes\infty}, \mu)$ into $(\widehat{X^{\otimes\infty}}, \hat\mu)$.

The Rokhlin-Abramov formula gives

$$h(\mu) = h(\hat\mu) + h(\mu \mid \pi_\mu),$$

where

$$h(\mu \mid \pi_\mu) = \int_{\widehat{X^{\otimes\infty}}} \lim_{n\to\infty} \frac{1}{|F_n|} H_\mu(F_n \mid \pi_\mu(x))\, d\mu$$

is the relative entropy of μ over π_μ. However, by Lemma 3.4, $h(\hat\mu) = 0$. Since $H_\mu(F_n \mid \pi_\mu(x)) \le S(\pi_\mu(x)_{F_n})$,

$$h_\infty(X) = h(\mu) = h(\mu \mid \pi_\mu) \le \int_{\widehat{X^{\otimes\infty}}} \limsup_{n\to\infty} \frac{1}{|F_n|} . S(\hat x_{F_n})\, d\hat\mu(\hat x) \tag{3}$$

However, for any $\hat x \in \widehat{X^{\otimes\infty}}$,

$$\limsup_{n\to\infty} \frac{1}{|F_n|} S(\hat x_{F_n}) \le h_{ind}(X),$$

so $h_\infty(X) \le h_{ind}(X)$. Finally, we already know from [15] that $h_{ind}(X) \le h_\infty(X)$, so we have shown that $h_\infty(X) = h_{ind}(X)$.

□

4. Limiting measures of maximal entropy

We now wish to discuss the structure of the measure(s) of maximal entropy μ on $X^{\otimes\infty}$. We will show that we can completely describe the structure of the measures of maximal entropy which are $\mathbb{P}$-invariant. Denote by $\widehat{X^{\otimes\infty}}_{\max} \subset \widehat{X^{\otimes\infty}}$ the set of points $\hat x$ with $S(\hat x) = h_{ind}(X)$. This is a shift-invariant and $\mathbb{P}$-invariant subset of $\widehat{X^{\otimes\infty}}$.

For any shift-invariant measure ν supported on $\widehat{X^{\otimes\infty}}_{\max}$, define a shift-invariant measure $\Phi(\nu)$ on $X^{\otimes\infty}$ by

$$\Phi(\nu) = \int_{\widehat{X^{\otimes\infty}}} \mu_{\hat x}\, d\nu(\hat x), \tag{4}$$

where for any $\hat{x} \in \widehat{X^{\otimes\infty}}_{\max}$, $\mu_{\hat{x}}$ is the independent product of the uniform measures over $\hat{x}_g$ for $g \in \mathbb{Z}^\infty$.

Theorem 4.1. *For any $\mathbb{P}$-invariant measure of maximal entropy μ on $X^{\otimes\infty}$, there exists a measure $\hat{\mu}$ supported on $\widehat{X^{\otimes\infty}}_{\max}$ such that $\mu = \Phi(\hat{\mu})$.*

Proof. If μ is a $\mathbb{P}$-invariant measure of maximal entropy for $X^{\otimes\infty}$, define $\hat{\mu} = \pi_\mu(\mu)$. Then,

$$h_\infty(X) = h_\mu(\{0\} \mid \pi_{P_\infty}) = \int_{X^{\otimes\infty}} H_\mu(\{0\} \mid \pi_{P_\infty})\, d\mu(x) = \int_{X^{\otimes\infty}} H_\mu(\{0\} \mid \pi_{P_\infty} \times \pi_\mu)\, d\mu(x)$$

$$\leq \int_{X^{\otimes\infty}} H_\mu(\{0\} \mid \pi_\mu)\, d\mu(x) \leq \int_{X^{\otimes\infty}} \log|\pi_\mu(x)_0|\, d\mu(x) = \int_{\widehat{X^{\otimes\infty}}} S(\hat{x}_0)\, d\hat{\mu}(\hat{x}) \leq h_{ind}(X).$$

The third equality holds since $h(\hat{\mu}) = 0$. By Theorem 1.1, $h_\infty(X) = h_{ind}(X)$. Therefore, all inequalities above are in fact equalities. The first inequality being an equality implies that x_0 is conditionally μ-independent from x_P, given $\pi_\mu(x)$. This clearly implies that μ-almost every fiber $\mu_{\hat{x}}$ in the disintegration of μ over π_μ is a sitewise independent product. The second inequality being an equality implies that μ-a.s., the distribution of $\mu_{\hat{x}}$ on a site $g \in \mathbb{Z}^\infty$ is uniform over $\hat{x}_g$. Finally, the third inequality being an equality implies that $\hat{\mu}$-a.s., $S(\hat{x}) = h_{ind}(X)$, and so $\hat{\mu}$ is supported on $\widehat{X^{\otimes\infty}}_{\max}$.

□

Theorem 4.2. *Φ is an injective map which sends shift-invariant measures supported on $\widehat{X^{\otimes\infty}}_{\max}$ to measures of maximal entropy on $X^{\otimes\infty}$. Also, Φ preserves $\mathbb{P}$-invariance of measures.*

Proof. Injectivity of Φ follows from the easily checked fact that $\pi_{\Phi(\hat{\mu})} \circ \Phi(\hat{\mu}) = \hat{\mu}$ for any measure $\hat{\mu}$ supported on $\widehat{X^{\otimes\infty}}_{\max}$.

For any such $\hat{\mu}$, define $\mu := \Phi(\hat{\mu})$. Then $h(\mu) = h(\mu \mid \pi_\mu) = \int_{\widehat{X^{\otimes\infty}}} \log|\hat{x}_0|\, d\hat{\mu}(\hat{x})$, but since $\hat{\mu}$ is supported on $\widehat{X^{\otimes\infty}}_{\max}$, $\int_{\widehat{X^{\otimes\infty}}} \log|\hat{x}_0|\, d\hat{\mu}(\hat{x}) = h_{ind}(X^{\otimes\infty})$. Thus, $h(\mu) = h_{ind}(X^{\otimes\infty})$,which is equal to $h(X^{\otimes\infty})$ by Theorem 1.1, so μ is a measure of maximal entropy.

If $\hat{\mu}$ is $\mathbb{P}$-invariant, and $\rho \in \mathbb{P}$, then

$$\mu \circ \rho = \int_{\widehat{X^{\otimes\infty}}} \mu_{\hat{x}}\, d\nu(\rho\hat{x}) = \int_{\widehat{X^{\otimes\infty}}} \mu_{\hat{x}}\, d\nu(\hat{x}) = \mu,$$

so μ is $\mathbb{P}$-invariant.

□

Theorems 4.1 and 4.2 show that there is a bijective correspondence between the $\mathbb{P}$-invariant measures of maximal entropy on $X^{\otimes\infty}$ and the $\mathbb{P}$-invariant measures supported on $\widehat{X^{\otimes\infty}}_{\max}$.

We now recall some technical facts, shown in [15], about independence entropy for $\mathbb{Z}$ SFTs. It was shown in Theorem 2 from [15] that for any $\mathbb{Z}$ SFT X, there exists a word $\hat{w} \in \mathcal{L}(\hat{X}, [1,m])$ such that $\hat{w}^\infty \in \hat{X}$ and $\frac{S(\hat{w})}{m} = h_{ind}(X)$. Any such $\hat{w}^\infty$ is called a *maximizing cycle* for $\hat{X}$. A word $\hat{w} \in \mathcal{L}(\hat{X})$ is called a *maximizing word* for $\hat{X}$ if $\hat{w}^\infty$ is a maximizing cycle for $\hat{X}$.

For completeness, we give a self-contained proof of the following refinement of the above statement, which is essentially in [15].

Lemma 4.3. *For any k-step $\mathbb{Z}$-SFT X, $\hat{X}$ has a maximizing word $\hat{w}$ with no repeated k-letter subword.*

Proof. Consider a maximizing word $\hat{w}$ for $\hat{X}$, and denote its length by m. If $\hat{w}$ has no repeated k-letter subword, we are done. Suppose then that $\hat{w}$ does have a repeated k-letter subword, call it u. Let's say that $\hat{w}_{[i,i+k-1]} = \hat{w}_{[j,j+k-1]} = u$, $i < j$. Define $\hat{a} = \hat{w}_{[i,j-1]}$ and $\hat{b} = \hat{w}_{[j,m]}\hat{w}_{[1,i-1]}$. We now claim that both $\hat{a}$ and $\hat{b}$ are also maximizing words for $\hat{X}$. Note that every k-letter subword of $\hat{a}^\infty$ was already a subword of $\hat{w}_{[i,j+k-1]}$, which is clearly in $\mathcal{L}(\hat{X})$ since $\hat{w}$ is. Similarly, every k-letter subword of $\hat{b}^\infty$ was already a subword of $\hat{w}_{[k,m]}\hat{w}_{[1,i+k-1]}$, which is in $\mathcal{L}(\hat{X})$ since $\hat{w}^2$ is. Therefore, both $\hat{a}^\infty$ and $\hat{b}^\infty$ are in $\hat{X}$.

Finally, we note that since each letter of $\hat{w}$ is contained in exactly one of $\hat{a}$ or $\hat{b}$, $S(\hat{w}) = h_{ind}(X)$ is a weighted average of $S(\hat{a})$ and $S(\hat{b})$. Both $S(\hat{a}) = S(\hat{a}^\infty)$ and $S(\hat{b}) = S(\hat{b}^\infty)$ are less than or equal to $h_{ind}(X)$ by definition, so both are equal to $h_{ind}(X)$, and therefore both $\hat{a}^\infty$ and $\hat{b}^\infty$ are maximizing cycles for $\hat{X}$. Then $\hat{a}$ and $\hat{b}$ are maximizing words for $\hat{X}$, and since each has length less than $\hat{w}$, we can continue this procedure until we arrive at a maximizing word for $\hat{X}$ with no repeated k-letter subwords. $\square$

We say that a maximizing word for a k-step $\mathbb{Z}$ SFT is *simple* if it has no repeated k-letter subwords, and that a maximizing cycle is simple if it can be written as $\hat{w}^\infty$ for a simple maximizing word $\hat{w}$. Now, for any simple maximizing word $\hat{w}$, we will construct a specific finite orbit contained in $\widehat{X^{\otimes\infty}}_{\max}$. The method is simple: define $x(\hat{w}) \in \hat{\Sigma}^{\mathbb{Z}^\infty}$ by $x(\hat{w})_g = \hat{w}_{\sum g_i \pmod{|\hat{w}|}}$. Then for any d, m and $g \in \mathbb{Z}^\infty$, $x(\hat{w})_{g+me_d} = \hat{w}_{\sum g_i + m \pmod{|\hat{w}|}}$, and so $x(\hat{w})_{g+\mathbb{Z}e_d}$ is just a shift of the sequence $\hat{w}^\infty$. Clearly, this implies that $x(\hat{w}) \in \widehat{X^{\otimes\infty}}_{\max}$. Denote by $\mathcal{O}(\hat{w})$ the finite ($\mathbb{P}$-invariant) orbit of $x(\hat{w})$ under $\mathbb{Z}^\infty$, and by $\hat{\mu}_{\hat{w}}$ the uniform measure on $\mathcal{O}(\hat{w})$. Then by Theorem 4.2, $\Phi(\hat{\mu}_{\hat{w}})$ is a $\mathbb{P}$-invariant measure of maximal entropy on $X^{\otimes\infty}$.

Theorem 4.4. *For any $\mathbb{Z}$ SFT X, there is a unique $\mathbb{P}$-invariant measure of maximal entropy on $X^{\otimes\infty}$ if and only if the following two conditions are satisfied:*

(1) *$\hat{X}$ has a unique (up to shifts) simple maximizing cycle $\hat{w}$*
(2) *There is only one finite orbit of points in $\hat{\Sigma}^{\mathbb{Z}^2}$ for which each row and each column is a shift of the sequence $\hat{w}^{\mathbb{Z}}$, namely the orbit of the periodic point $\hat{w}^{(2)}$ defined by $\hat{w}^{(2)}_{i,j} = \hat{w}_{i+j \pmod{|\hat{w}|}}$.*

Proof. ($\Longrightarrow$) If condition (1) is violated, then $\hat{X}$ has two simple maximizing cycles $\hat{w}^\infty$ and $\widehat{w'}^\infty$ which are not shifts of each other, which induce points $x(\hat{w})$ and $x(\widehat{w'})$ with different finite orbits $\mathcal{O}(\hat{w})$ and $\mathcal{O}(\widehat{w'})$ contained in $\widehat{X^{\otimes\infty}}_{\max}$. Then by injectivity of Φ, $\Phi(\hat{\mu}_{\hat{w}})$ and $\Phi(\hat{\mu}_{\widehat{w'}})$ are distinct $\mathbb{P}$-invariant measures of maximal entropy on $X^{\otimes\infty}$.

If condition (1) is satisfied but condition (2) is violated, then $\hat{X}$ has a unique (up to shifts) simple maximizing cycle $\hat{w}^\infty$ and a point $\widehat{w'}^{(2)}$ in $\hat{\Sigma}^{\mathbb{Z}^2}$ whose rows and

columns are all shifts of $\hat{w}^\infty$, but which is not a shift of $\hat{w}^{(2)}$. We now construct an uncountable family of shift-invariant and $\mathbb{P}$-invariant measures supported on $\widehat{X^{\otimes\infty}}_{\max}$, which by Theorem 4.2 will yield an uncountable family of $\mathbb{P}$-invariant measures of maximal entropy on $X^{\otimes\infty}$. First, for any $\alpha \in (0, 0.5)$, define η_α to be the Bernoulli measure on $\{0,1\}^{\mathbb{N}}$ which gives probability α to 0 and $1-\alpha$ to 1 at each site. Define υ to be the uniform measure on $\{0,1,\ldots,|\hat{w}|-1\}$. Define a factor map τ from $\{0,1,\ldots,|\hat{w}|-1\}^2 \times \{0,1\}^{\mathbb{N}}$ to $\hat{\Sigma}^{\mathbb{Z}^\infty}$ by

$$\tau(i,j,(u_n))(g) := \widehat{w'}^{(2)}\left(i + \sum_{\{n \,:\, u_n=0\}} g_n,\ j + \sum_{\{n \,:\, u_n=1\}} g_n\right)$$

We first claim that τ maps to $\widehat{X^{\otimes\infty}}_{\max}$. This is easy to check: by definition, for any i, j, and (u_n), every row of $\tau(i,j,(u_n))$ is just a row or column of $\widehat{w'}^{(2)}$, which will always be a shift of $\hat{w}^{\mathbb{Z}}$. Now, for any $\alpha \in (0,0.5)$, define μ_α to be the push-forward of $\upsilon \times \upsilon \times \eta_\alpha$ under τ. Clearly each μ_α is a measure on $\widehat{X^{\otimes\infty}}_{\max}$. The reader may check that $\mathbb{P}$-invariance of μ_α follows from the fact that η_α is i.i.d., and that shift-invariance of μ_α follows from the uniformity of υ and the fact that $\widehat{w'}^{(2)}$ is periodic with respect to $(|\hat{w}|, 0)$ and $(0, |\hat{w}|)$. All that remains is to show that all μ_α are distinct. For any α, define ν_α to be the marginalization of μ_α to $\mathbb{Z}^2 \times \{0\}^\infty$. It is clear that ν_α is always a finitely supported measure, which is a shift of $\hat{w}^{(2)}$ with probability $\alpha^2 + (1-\alpha)^2$, a shift of $\widehat{w'}^{(2)}$ with probability $\alpha(1-\alpha)$, and a shift of $\underline{\widehat{w'}^{(2)}}$ with probability $\alpha(1-\alpha)$, where $\underline{\widehat{w'}^{(2)}}$ is obtained from $\widehat{w'}^{(2)}$ by permuting the first and second coordinates. But then since $\hat{w}^{(2)}$ is different from both $\widehat{w'}^{(2)}$ and $\underline{\widehat{w'}^{(2)}}$, and since $\alpha^2 + (1-\alpha)^2$ is injective on $(0,0.5)$, clearly all ν_α are distinct, implying that all μ_α are distinct, and therefore that all $\Phi(\mu_\alpha)$ are distinct $\mathbb{P}$-invariant measures of maximal entropy on $X^{\otimes\infty}$.

($\Longleftarrow$) If conditions (1) and (2) are satisfied, then we will show that any shift-invariant measure supported on $\widehat{X^{\otimes\infty}}_{\max}$ is in fact supported on $\mathcal{O}(\hat{w})$. Clearly the only such measure is $\hat{\mu}_{\hat{w}}$, and Theorem 4.1 then implies that the only $\mathbb{P}$-invariant measure of maximal entropy on $X^{\otimes\infty}$ is $\Phi(\hat{\mu}_{\hat{w}})$.

Suppose that $\hat{\mu}$ is a shift-invariant measure on $\widehat{X^{\otimes\infty}}_{\max}$, and choose any $d \in \mathbb{N}$. Then $S(\hat{\mu}) = \int_{\widehat{X^{\otimes\infty}}_{\max}} S(\hat{x}_{\mathbb{Z}e_d})\, d\hat{\mu}(\hat{x}) = h_{ind}(X)$, so clearly $S(\hat{x}_{\mathbb{Z}e_d}) = h_{ind}(X)$ $\hat{\mu}$-almost surely. Choose k so that $\hat{X}$ is a k-step SFT, and consider any $\hat{u}$ which contains no repeated k-letter word and for which $\hat{u}^\infty \in \hat{X}$ is not a shift of $\hat{w}^\infty$. Since $\hat{w}^\infty$ was the unique (up to shifts) simple maximizing cycle for $\hat{X}$, $S(\hat{u}) < h_{ind}(X)$. If $\hat{\mu}([\hat{u}]) > 0$, then for $\hat{\mu}$-a.e. $\hat{x} \in \hat{X}$, $\hat{u}$ occurs within $\hat{x}_{\mathbb{Z}e_d}$ with positive frequency. Denote by $\hat{v}$ the k-letter prefix of $\hat{u}$. Then we can decompose any $\hat{x}_{\mathbb{Z}e_d}$ as $\ldots \hat{v}u_{-1}\hat{v}u_0\hat{v}u_1\hat{v}\ldots$, where $\hat{v}u_i = \hat{u}$ for a set of integers i of positive density. Then, since $(\hat{v}u_i)^\infty \in \hat{X}$ for every i, the same argument from Lemma 4.3 shows that $S(\hat{v}u_i) \leq h_{ind}(X)$ for all i, and so $S(\hat{x}_{\mathbb{Z}e_d}) < h_{ind}(X)$, a contradiction. So, $\hat{\mu}([\hat{u}]) = 0$. Clearly, this shows that $\hat{\mu}$-a.s., $\hat{x}_{\mathbb{Z}e_d}$ is just a shift of $\hat{w}^\infty$. By shift-invariance, for $\hat{\mu}$-a.e. $\hat{x} \in \widehat{X^{\otimes\infty}}$ it is also the case that for any $g \in \mathbb{Z}^\infty$ and $d \in \mathbb{N}$, $\hat{x}_{g+\mathbb{Z}e_d}$ is also a shift of $\hat{w}^\infty$.

Consider any such $\hat{x}$, where every row in every direction is a shift of $\hat{w}^\infty$. Then for any dimensions $d_1 < d_2$ and any $g \in \mathbb{Z}^\infty$, consider the infinite two-dimensional configuration $\hat{x}_{g+\mathbb{Z}e_{d_1}+\mathbb{Z}e_{d_2}}$. Each row and column of $\hat{x}_{g+\mathbb{Z}e_{d_1}+\mathbb{Z}e_{d_2}}$ is a shift of $\hat{w}^\infty$. But then by condition (2), $\hat{x}_{g+\mathbb{Z}e_{d_1}+\mathbb{Z}e_{d_2}}$ is a shift of $\hat{w}^{(2)}$ and so is periodic with respect to $e_{d_1} - e_{d_2}$. Since this is true for all d_1, d_2, and g, $\hat{x}$ must be periodic with respect to $e_{d_1} - e_{d_2}$ for all $d_1 < d_2$. This implies in turn that $\hat{x}$ is periodic with respect to any $g \in \mathbb{Z}^\infty$ with $\sum g_i = 0$. It is simple to check that any point in $\hat{\Sigma}^{\mathbb{Z}^\infty}$ which is periodic with respect to all such vectors and whose marginalization to $\mathbb{Z}e_1$ is a shift of $\hat{w}^\infty$ must be a shift of $x(\hat{w})$.

We have then shown that any shift-invariant measure $\hat{\mu}$ on $\widehat{X^{\otimes\infty}}_{\max}$ is supported on $\mathcal{O}(x(\hat{w}))$, which implies that $\Phi(\hat{\mu}_{\hat{w}})$ is the unique $\mathbb{P}$-invariant measure of maximal entropy on $X^{\otimes\infty}$ as explained above.

□

The techniques of Theorem 4.4 also allow us to give one more case in which the set of $\mathbb{P}$-invariant measures of maximal entropy can be completely described.

Theorem 4.5. *If X is a $\mathbb{Z}$ SFT such that $\hat{X}$ has k different (up to shifts) simple maximizing cycles $\widehat{w_i}$, $1 \leq i \leq k$, and if no two $\widehat{w_i}$ share a common letter of $\hat{\Sigma}$, and if for each $i \in [1,k]$, there is only one finite orbit of points in $\hat{\Sigma}^{\mathbb{Z}^2}$ for which each row and each column is a shift of the sequence $\widehat{w_i}^{\mathbb{Z}}$, then there are exactly k ergodic $\mathbb{P}$-invariant measures of maximal entropy on $X^{\otimes\infty}$.*

Proof. We will only sketch a proof, as the details are almost the same as in the proof of Theorem 4.4. Firstly, by the same reasoning used there, for any ergodic $\mathbb{P}$-invariant measure of maximal entropy μ on $X^{\otimes\infty}$ and for any $g \in \mathbb{Z}^\infty$ and $d \in \mathbb{N}$, it is $\hat{\mu}$-a.s. the case that $\hat{x}_{g+\mathbb{Z}e_d}$ is a shift of one of the simple maximizing cycles $\widehat{w_i}^\infty$. But then since no two $\widehat{w_i}$ share a common letter, this i must be the same for all g and d, and by ergodicity, it is $\hat{\mu}$-a.s. constant. So, there exists i for which $\hat{\mu}$-a.s., for any $g \in \mathbb{Z}^\infty$ and $d \in \mathbb{N}$, $\hat{x}_{g+\mathbb{Z}e_d}$ is a shift of $\widehat{w_i}^\infty$.

Then a similar argument as was used above shows that $\hat{\mu}$ must be supported on $\mathcal{O}(x(\widehat{w_i}))$, and so $\mu = \Phi(\hat{\mu}_{\widehat{w_i}})$. Since there are only k possible choices for i, and each clearly gives a different measure, we are done.

□

5. Applications to specific models

The purpose of this section is present some applications of our general results to various specific models which have appeared in the litrature.

5.1. **Hard-square model.**

The underlying $\mathbb{Z}$ subshift, also known as the golden mean shift, is

$$\mathcal{H} := \{x \in \{0,1\}^{\mathbb{Z}} \ : \ x_n x_{n+1} \neq 11\}.$$

The d-dimensional hard-square model is then defined as $\mathcal{H}^{\otimes d}$.

It is easily checked that $h_\infty(\mathcal{H}) = h_{ind}(\mathcal{H}) = \frac{1}{2}\log(2)$. As shown in [15], the results of Galvin and Kahn [7] can be used to show that $h(\mathcal{H}^{\otimes d}) \to h_\infty(\mathcal{H})$ at an exponential rate, with explicit numerical bounds. In this case, it is easily checked that $(\overline{0\{0,1\}})^\infty$ is the unique (up to shifts) simple maximizing cycle for $\hat{\mathcal{H}}$, and so Theorem 4.4 implies that there is a unique $\mathbb{P}$-invariant measure of maximal entropy on $\mathcal{H}^{\otimes\infty}$. In fact by [7], uniqueness holds even without the assumption of

$\mathbb{P}$-invariance. The unique measure of maximal entropy is not weak mixing; since μ-a.s. each point of $\mathcal{H}^{\otimes\infty}$ either has 0s on all odd sites or 0s on all even sites (the parity of $v \in \mathbb{Z}^\infty$ is just the parity of the sum of its coordinates), and μ-a.s. only one of these can hold, clearly μ has an eigenfunction with eigenvalue of -1. The combinatorial methods of Galvin and Kahn show that this eigenvalue is also present in the (unique) measure of maximal entropy on $\mathcal{H}^{\otimes d}$ for all sufficiently large d.

5.2. **n-coloring shifts.** The one-dimensional n-coloring shift is

$$\mathcal{C}_n := \{x \in \{1, \dots, n\}^{\mathbb{Z}} \ : \ x_n \neq x_{n+1}\}.$$

The d-dimensional n-coloring shift is defined as $\mathcal{C}_n^{\otimes d}$.

By our results, $h_\infty(\mathcal{C}_n) = h_{ind}(\mathcal{C}_n) = \frac{1}{2}\log(\lfloor n/2 \rfloor) + \frac{1}{2}\log(\lceil n/2 \rceil)$. There are $\binom{n}{\lfloor n/2 \rfloor}$ (up to shifts) simple maximizing cycles for $\hat{\mathcal{C}_n}$, namely all sequences $(\hat{a}\hat{b})^\infty$ for which $\hat{a}$ and $\hat{b}$ form a partition of Σ and $|\hat{a}| = \lfloor \frac{|\Sigma|}{2} \rfloor$. Since no two of these cycles share a common letter, by Theorem 4.5 there are exactly $\binom{n}{\lfloor n/2 \rfloor}$ ergodic $\mathbb{P}$-invariant measures of maximal entropy on $h(\mathcal{C}_n^{\otimes\infty})$, each of which has eigenvalue -1 as in the hard-square model.

The case $n = 3$ is of particular interest, since 3-colorings have useful connections to the well known "square-ice" model for $d = 2$ [13]. In [15], it was shown that $h(\mathcal{C}_3^{\otimes d}) \to h_\infty(\mathcal{C}_3) = \frac{\ln 2}{2}$ exponentially fast. The argument involved creating a correspondence between configurations in the d-dimensional hard-square and 3-coloring shifts, and then exploiting the previously mentioned results of Galvin and Kahn. From a recent paper of Peled [20], it follows that for sufficiently large d there are exactly 3 measures of maximal entropy (even without the assumption of $\mathbb{P}$-invariance), each of which admits an eigenfunction with eigenvalue -1, just as in the hard-square case.

5.3. **Beach model.** In [2], Burton and Steif defined the *d-dimensional beach model*, for any $M > 0$, to be the nearest-neighbor $\mathbb{Z}^d$ SFT on the alphabet $\{-M, \dots, -1, 1, \dots, M\}$ defined by the restriction that adjacent letters must have product greater than or equal to -1. In other words, a negative and positive cannot be adjacent in any cardinal direction unless they are 1 and -1. These are clearly all axial powers of the one-dimensional beach model, which we denote by B_M.

It is easy to show that $h_{ind}(B_M) = \log M$, and that for $M > 2$, there are exactly two simple maximizing cycles for B_M, namely $\{-M, \dots, -1\}^\infty$ and $\{1, \dots, M\}^\infty$. (When $M = 1$, B_M is just the full shift on two symbols, and when $M = 2$, there is an additional maximizing cycle $\{-1, 1\}^\infty$. We will not address these special cases further here.) Therefore, our results show that $h(B_M^{\otimes d}) \to \log M$, and that $B_M^{\otimes\infty}$ has exactly two ergodic $\mathbb{P}$-invariant measures of maximal entropy.

In fact, it was also shown in [2] that for any fixed d and for $M > 4e28^d$, $B_M^{\otimes d}$ has exactly two ergodic measures of maximal entropy. Our result would seem to suggest that the same is true for fixed M and sufficiently large d, and in fact this was stated as a conjecture in [3].

The beach model was further studied in [9], where it was extended to a more general class of models, with a real-valued parameter M replacing the previous integer-valued parameter, Disagreement percolation techniques were used to show that for any fixed d and $M < \frac{2d^2+d+1}{2d^2+d-1}$, the d-dimensional beach model has a unique measure of maximal entropy. Unfortunately, this only applies to values of M less than 2, and so never applies to the classical beach model.

5.4. **Run length limited shifts.** For any $0 \leq d < k \leq \infty$, the (d,k) *run-length limited shift*, also denoted by $RLL(d,k)$, is the SFT on the alphabet $\{0,1\}$ consisting of all sequences in which all maximal "runs" of 0s have length inside the interval $[d,k]$. For instance, $RLL(0,\infty)$ is the full shift on two symbols, and $RLL(1,\infty)$ is the usual golden mean shift. For any $0 \leq d < k < \infty$, it was shown in [21] that

$$h_{ind}(RLL(d,k)) = \frac{\lfloor (k-d)/(d+1) \rfloor \ln 2}{\lfloor (k+1)/(d+1) \rfloor (d+1)} \text{ and}$$

$$h_{ind}(RLL(d,\infty)) = \frac{\ln 2}{d+1}.$$

It was shown in [18], using combinatorial methods, that for any d, $h(RLL(d,\infty)^{\otimes d}) \to h_{ind}(RLL(d,\infty))$, and that the rate is exponential. Our results show that this same convergence is true for any d and k (but say nothing about the rate.)

It is relatively simple to check that there is a unique (up to shifts) simple maximizing cycle for any d and k. (This was essentially done, without using our terminology, in [21].) These cycles are given by the maximizing words

$$\hat{w} = 0^d\{0,1\} \text{ for } RLL(d,\infty) \text{ and}$$

$$\hat{w} = 0^d 1 (0^d\{0,1\})^{\lfloor (k-d)/(d+1) \rfloor} \text{ for } RLL(d,k).$$

All of these unique maximizing cycles $\hat{w}^\infty$ are symmetric, meaning that $(\hat{w}_{|\hat{w}|}\hat{w}_{|\hat{w}|-1}\dots\hat{w}_1)^\infty$ is just a shift of $\hat{w}^\infty$. Therefore, for any $RLL(d,k)$, we can construct a point $\widehat{w'}^{(2)}$ defined by $\widehat{w'}^{(2)}_{i,j} = \hat{w}_{i-j \pmod{|\hat{w}|}}$, in which all rows and columns are shifts of $\hat{w}^\infty$. In all cases except $RLL(1,\infty)$ and $RLL(0,1)$ (for which the associated simple maximizing words have length 2) and $RLL(0,\infty)$ (for which the associated simple maximizing word has length 1), $\widehat{w'}^{(2)}$ is not equal to the point $\hat{w}^{(2)}$ from Lemma 4.4, and so if X is any run-length limited shift except $RLL(1,\infty)$, $RLL(0,1)$, or $RLL(0,\infty)$, then $X^{\otimes\infty}$ has multiple $\mathbb{P}$-invariant measures of maximal entropy. For each of these three special cases, $X^{\otimes\infty}$ has a unique $\mathbb{P}$-invariant measure of maximal entropy, which we already knew, as these are simply the golden mean shift, the golden mean shift with digits 0 and 1 reversed, and the full shift on two symbols, respectively.

5.5. **Even shift.** The even shift is the sofic shift $\mathcal{E}$ on the alphabet $\{0,1\}$ consisting of all sequences in which all maximal "runs" of 0s have even length. It is easy to show that $h_{ind}(\mathcal{E}) = 0$; in fact, $\hat{\Sigma} = \{\{0\},\{1\},\{0,1\}\}$, and it is not hard to check that the maximum number of times the symbol $\{0,1\}$ can appear in a point of $\hat{\mathcal{E}}$ is two. So, our results imply that $h_\infty(\mathcal{E}) = 0$.

Finding $h_\infty(\mathcal{E})$ was of particular interest for two reasons. Firstly, in [14], a combinatorial argument was used to show that for the similarly defined odd shift $\mathcal{O}$, $h(\mathcal{O}^{\otimes d}) = \frac{1}{2}$ for all d. Secondly, it was shown in [15] that $h_\infty(X) = 0$ for any $\mathbb{Z}$ SFT X with zero independence entropy, and it was naturally wondered if the same was true for sofic shifts.

5.6. **Dyck shift.** This non-sofic subshift $\mathcal{D} \subset (\{\alpha_1,\dots,\alpha_M\} \cup \{\beta_1,\dots,\beta_M\})^{\mathbb{Z}}$ is obtained by considering the alphabet letters as M "types" of "brackets." The constraints are that matching open and closed brackets must be of the same "type," which in our terminology means the same subscript. This interesting non-sofic shift originated in the study of formal languages. It was introduced into symbolic

dynamics in [12], where it was shown that $h(\mathcal{D}) = \log(M+1)$ and that there are exactly 2 ergodic measures of maximal entropy.

It is easily shown that $h_{ind}(\mathcal{D}) = \log M$, and that there are precisely two (up to shifts) simple maximizing cycles on $\hat{\mathcal{D}}$, namely $\{\alpha_1, \ldots, \alpha_M\}^\infty$ and $\{\beta_1, \ldots, \beta_M\}^\infty$. Thus $h_\infty(\mathcal{D}) = \log M$, and since these cycles contain no common letter, Theorem 4.5 shows that there are precisely two ergodic $\mathbb{P}$-invariant measures of maximal entropy on $\mathcal{D}^{\otimes\infty}$.

5.7. **Symmetric nearest-neighbor SFTs.** In recent work of Engbers and Galvin [5], they study, for any finite undirected graph $\mathbb{H}$, the limiting behavior of the distribution of uniformly chosen graph homomorphisms from discrete d-dimensional m-tori $\mathbb{Z}_m^d = \{1, \ldots, m\}^d$ to $\mathbb{H}$ as $d \to \infty$. Recall that a *graph homomorphism* from $G_1 = (V_1, E_1)$ to $G_2 = (V_2, E_2)$ is a function $x : V_1 \to V_2$ such that $(x(v), x(w))$ is an edge in E_2 whenever $(v, w) \in E_1$. In particular, for a fixed finite undirected graph $G = (V, E)$, the collection of graph homomorphisms from $\mathbb{Z}^d$ to G is precisely the nearest-neighbor SFT $X_G \subset V^{\mathbb{Z}^d}$ defined by enforcing $(x_n, x_{n+e_i}) \in E$ for all $n \in \mathbb{Z}^d$ and $i = 1 \ldots d$, which is a d-dimensional axial power of a symmetric nearest-neighbor SFT.

The authors prove that for any fixed m and undirected finite graph $\mathbb{H} = (\mathbb{V}, \mathbb{E})$, with probability tending to 1 as $d \to \infty$, a uniformly chosen random graph homomorphism x from $\mathbb{Z}_m^d$ to $\mathbb{H}$ has corresponding disjoint $A, B \subset \mathbb{V}$ which induce a complete bipartite graph, with $|A||B|$ maximal, such that $x_n \in A$ for a large proportion of even vertices $n \in \mathbb{Z}_m^d$ and $x_m \in B$ for a large proportion of odd vertices $m \in \mathbb{Z}_m^d$.

These results are related to ours in that they provide an alternative proof for some of our results for the particular case of symmetric nearest-neighbor SFTs. Observe that for any symmetric nearest-neighbor SFT $X \subset \Sigma^{\mathbb{Z}}$, there is always a simple maximizing word in $\widehat{X}$ of length 2, and that any such word corresponds to a maximal induced complete bipartite graph of $\mathbb{H}$.

However, the methods of Engbers and Galvin are apparently very different then ours, and in particular give additional finitistic results.

6. Further Problems and research directions

Here we summarize a few possible directions for extensions or generalizations of our results.

6.1. **Pressure and equilibrium measures.** In statistical mechanics, it is common to introduce a "potential" or "activity function," in which case the role of topological entropy is replaced by topological pressure, and measures of maximal entropy are generalized to equilibrium measures. From the ergodic-theoretic point of view, many results generalize without difficulty (see, for example, [22] for a statement and proof of the variational principle for pressure).

In [7] and in [5], for some specific $\mathbb{Z}$ subshifts, an analysis of equilibrium measures was carried out with respect to a single-site potential on $X^{\otimes d}$ as $d \to \infty$.

It is rather easy to generalize the statements and proofs of our results to the setting where entropy is replaced by pressure with respect to a so-called "single-site potential" $f : X \to \mathbb{R}$, given by $f(x) = F(x_0)$ for some function $F : \Sigma \to \mathbb{R}$. From the statistical mechanics viewpoint, f doesn't involve any "interactions" between sites.

Here is a brief formulation of the analogous result: For $\hat{w} \in \mathcal{L}(\hat{X}, F)$, define $S_f(\hat{w}) = \dfrac{1}{F} \sum_{n \in F} \log \sum_{a \in \hat{w}_n} F(a)$. Define $P_{ind}(X, f)$ analogously to $h_{ind}(X)$, with S_f replacing S, and

$$P_\infty(X, f) = \lim_{d \to \infty} P(X^{\otimes d}, f^{(d)}),$$

where $f^{(d)} : X^{\otimes d} \to \mathbb{R}$ is again given by $f(x) = F(x_0)$.

By following our proof of Theorem 1.1, one can deduce that $P_{ind}(X, f) = P_\infty(X, f)$.

It is more difficult to generalize to interactions involving multiple sites, even to the case where f depends only on the letters at the pair of adjacent sites x_0 and x_1. We believe a version of our result to be true for any interaction f which is finite-range (meaning that f depends only on the letters at finitely many sites), but it is not completely clear what the proper hypotheses would be, and in particular what a "natural" definition of $f^{(d)}$ would be in this case.

6.2. **Finitistic results.** Our techniques involve studying the system $X^{\otimes \infty}$, which is a sort of "infinite-dimensional" axial power of X. It is natural to wonder what information we can glean about the finite-dimensional axial powers $X^{\otimes d}$. For instance, we have shown that $h(X^{\otimes d}) \to h_{ind}(X)$, but with no information about the rate. This is a question of particular interest, since it was noted in [15] that for all examples where the rate of convergence is known, this rate is exponential.

Another example of useful finite-dimensional information regards measures of maximal entropy. As described in Section 5, our results allow us to count the $\mathbb{P}$-invariant measures of maximal entropy on $X^{\otimes \infty}$ for many models, such as the hard-square model and n-coloring shift. It is natural to assume that such results should allow us to draw conclusions about the number of $\mathbb{P}$-invariant measures of maximal entropy on $X^{\otimes d}$ for large enough d, but we have not yet been able to do so.

One reason why we believe such finitistic results should be provable is that one of the keys to our proofs, de Finetti's Theorem, has versions which apply to finite sets of exchangeable random variables. ([4]) We hope to use these finite versions to answer some finitistic questions in future work.

One specific case in which exponential convergence of $h(X^{\otimes d})$ to $h_{ind}(X)$ has been proven is in the case where X is an SFT and $h_{ind}(X) = 0$. ([15]) Interestingly, in the case where X is nearest neighbor, this convergence is trivial.

Lemma 6.1. *For any nearest-neighbor $\mathbb{Z}$ SFT X with zero independence entropy, $h(X^{\otimes 2}) = 0$.*

Proof. Let $X \subset \Sigma^{\mathbb{Z}}$ be a n.n. SFT with zero independence entropy. We can restrict to the non-wandering part of X, which is a disjoint union of irreducible subshifts of finite type, thus we may assume X is irreducible without loss of generality.

First we show that for any $a, b \in \Sigma$, there is at most one $c \in \Sigma$ such that $acb \in \mathcal{L}(X)$. If this were not the case, there would exist distinct $c_1, c_2 \in \Sigma$ with $ac_ib \in \mathcal{L}(X)$ for $i = 1, 2$, and then for some $d_1, \ldots, d_k \in \Sigma$, $\{a\}\{c_1, c_2\}\{b\}\{d_1\} \cdots \{d_k\}\{a\}$ would be in $\mathcal{L}(\hat{X})$, which would yield a point $(\{a\}\{c_1, c_2\}\{b\}\{d_1\} \cdots \{d_k\})^\infty$ with positive independence score, contradicting $h_{ind}(X) = 0$.

Next we claim that for any finite $F \subset \mathbb{Z}$, and any $c \in \Sigma^{\partial F}$ there is at most one $X^{\otimes 2}$-admissible configuration $d \in A^{F \cup \partial F}$ with $d_{\partial F} = c$; this is done by induction on $|F|$.

The base case of the induction is $|F| = 1$; say $F = \{(n,m)\}$. Then $(n-1,m),(n,m+1) \in \partial F$. Let $a = c_{(n-1,m)}$ and $b = c_{(n,m+1)}$: it follows that there is at most one c for which $acb \in \mathcal{L}(X)$. Since X is nearest neighbor, this means that there is at most one c for which ac and cb are both X-admissible. Thus, $d_{(n,m)}c$ is the only $X^{\otimes 2}$-admissible filling.

For the inductive step, choose $(n,m) \in F$ so that $(n-1,m),(n,m+1) \in \partial F$. (For instance, take (n,m) to be the greatest element of F lexicographically.) By the same argument, there is at most one letter $d_{(n,m)}$ which can fill (n,m) in a $X^{\otimes 2}$-admissible way given $c_{(n-1,m)}$ and $c_{(n,m+1)}$. Now the induction hypothesis on $F \setminus \{(n,m)\}$ implies that there is at most one $X^{\otimes 2}$-admissible filling of $F \cup \partial F$ given c and $d_{(n,m)}$.

This implies that $|\mathcal{L}(X^{\otimes 2},[1,n]^2)| \leq |\Sigma|^{|\partial[1,n]^2|}$ for any n, and so that $h(X^{\otimes 2}) = 0$.

□

For subshifts of finite type which are not nearest-neighbor, there is no finite d for which zero independence entropy implies $h(X^{\otimes d}) = 0$; it was demonstrated in [11] that $h(RLL(n,k)^{\otimes d}) = 0$ iff $k = n+1$, yet $h_{ind}(RLL(n,k)) = 0$ if $k \leq 2n$.

References

[1] R. Bowen. Entropy-expansive maps. *Trans. Amer. Math. Soc.*, 164:323–331, 1972.

[2] R. Burton and J. Steif. Non-uniqueness of measures of maximal entropy for subshifts of finite type. *Ergodic Theory Dynam. Systems*, 14:213–335, 1994.

[3] R. Burton and J. Steif. Some 2-d symbolic dynamical systems: entropy and mixing. In *Ergodic Theory and $\mathbb{Z}^d$-actions*, volume 228 of *London Math. Soc. Lecture Notes*, pages 297–306. Cambridge University Press, Cambridge, 1996.

[4] P. Diaconis and D. Freedman. Finite exchangeable sequences. *Ann. Probab.*, 8(4):745–764, 1980.

[5] J. Engbers and D. Galvin. H-coloring tori. Technical Report arXiv:1101.0840, Jan 2011. Comments: 39 pages.

[6] H. Furstenberg. *Recurrence in ergodic theory and combinatorial number theory.* Princeton University Press, Princeton, N.J., 1981. M. B. Porter Lectures.

[7] D. Galvin and J. Kahn. On phase transition in the hard-core model on $\mathbb{Z}^d$. *Combin. Probab. Comput.*, 13(2):137–164, 2004.

[8] H.-O. Georgii. *Gibbs measures and phase transitions*, volume 9 of *de Gruyter Studies in Mathematics.* Walter de Gruyter & Co.

[9] O. Häggström. On phase transitions for subshifts of finite type. *Israel J. Math.*, 94:319–352, 1996.

[10] E. Hewitt and L. J. Savage. Symmetric measures on Cartesian products. *Trans. Amer. Math. Soc.*, 80:470–501, 1955.

[11] H. Ito, A. Kato, Z. Nagy, and K. Zeger. Zero capacity region of multidimensional run length constraints. *Electron. J. Combin.*, 6:Research Paper 33, 16 pp. (electronic), 1999.

[12] W. Krieger. On the uniqueness of the equilibrium state. *Theory of Computing Sytstems*, 8:97–104, 1974.

[13] E. Lieb. Residual entropy of square ice. *Phys. Rev.*, 162(1):162–172, 1967.

[14] E. Louidor and B. Marcus. Improved lower bounds on capacities of symmetric 2d constraints using rayleigh quotients. *IEEE Trans. Inform. Theory*, 56(4):1624–1639, 2010.

[15] E. Louidor, B. Marcus, and R. Pavlov. Independence entropy of $\mathbb{Z}^d$-shift spaces. *Acta. Applicandae Mathematicae*, to appear.

[16] M. Misiurewicz. A short proof of the variational principle for a $\mathbf{Z}_+^N$ action on a compact space. In *International Conference on Dynamical Systems in Mathematical Physics (Rennes, 1975)*, pages 147–157. Astérisque, No. 40. Soc. Math. France, Paris, 1976.

[17] J. Moulin Ollagnier. *Ergodic theory and statistical mechanics*, volume 1115 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1985.

[18] E. Ordentlich and R. M. Roth. Independent sets in regular hypergraphs and multidimensional runlength-limited constraints. *SIAM J. Discrete Math.*, 17(4):615–623 (electronic), 2004.

[19] D. S. Ornstein and B. Weiss. Entropy and isomorphism theorems for actions of amenable groups. *J. Analyse Math.*, 48:1–141, 1987.

[20] R. Peled. High-dimensional lipschitz functions are typically flat. Technical Report arXiv:1005.4636v1, May 2010. Comments: 87 pages.

[21] T. L. Poo, P. Chaichanavong, and B. H. Marcus. Tradeoff functions for constrained systems with unconstrained positions. *IEEE Trans. Inform. Theory*, 52(4):1425–1449, 2006.

[22] P. Walters. A variational principle for the pressure of continuous transformations. *Amer. J. Math.*, 97(4):937–971, 1975.

[23] T. Ward and Q. Zhang. The Abramov-Rokhlin entropy addition formula for amenable group actions. *Monatsh. Math.*, 114(3-4):317–329, 1992.